Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 10 (2013) 978 – 986

Procedia Technology

International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013

Preservation of IPR of Songs through Embedding Secret Song (IPRSESS)

Uttam Kr. Mondal[a], J. K. Mandal[b,*]

[a]Dept. Of CSE & IT, College of Engg. & Management, Kolaghat, Midnapur (W.B), India
[b]Dept. Of CSE, University of Kalyani, Nadia (W.B), India

Abstract

Security is one of the major challenging issues for creative industries all over the world. This paper proposes a technique to provide security to digital song by embedding another strip of song as secret code into original song without affecting its audible quality. Separating phases and amplitudes of both song signals is precursor to the proposed technique followed by shifting of phases of secret song with modulating amplitude to fabricate the authenticating code. The embedded unique secret audio is used to detect and identify the originality of song from similar available songs. A comparative study has been made with existing techniques and experimental results are also presented based on Microsoft WAVE (".wav") stereo sound file.

Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keywords: Average absolute difference (AAD); maximum difference (MD); mean square error (MSE); normalized mean square error (NMSE); shifting of phases; song authentication; embedding secret song; IPR

1. Introduction

Today's creative organizations are facing competitive market for spreading business and holding market. Creating a quality product involved a lot of investment as well as effort. People are finding easier ways to put less effort or investing money and producing product for existence in this contemporary market. Some of them are

* J. K. Mandal. Tel.:+91 9434352214 , Fax: +91-33-25809617.
E-mail address: [2]jkm.cse@gmail.com

2212-0173
Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering
doi:10.1016/j.protcy.2013.12.446

applying technology for making piracy of original versions and producing lower price products. This intension is a frequent phenomenon for digital audio/video industries with improvement of digital editing technology [4, 8, 9, 10]. Even, it is difficult to listeners to find the original from pirated versions. Therefore, it is a big challenge for business persons, computer professionals or other concern people to ensure the security criteria of original songs and protect from releasing the duplicate versions.

In this paper, a framework for identifying a particular song with the help of unique secret message (concealment of another strip of song signal) over original song without altering its audible quality has been presented. The proposed technique is performed with sampled data (PCM format stereo type) by embedding short length song information with different components of original song in such a way that it will not alter the overall audible quality of the original song. The modified song can be separated easily from similar available songs using the secret embedded signal. It is experimentally observed that modified sampled values are not varied from the original audible quality.

Section 2.1 deals with the proposed embedding technique. The authentication procedure has been depicted in section 2.2. Experimental results are given in section 3. Conclusions are drawn in section 4. References are outlined at end.

2. Proposed Technique

The signal (song) is sampled and stored into an array x(n,2),where n is the number of rows, 2 is the number of columns representing the stereo channels. It is observed that the most of rows of the array having same values for two columns and it also observed that the remaining unequal values of sampled data having small difference. Experimentally it is seen that if we remove the extra magnitude values to make magnitude values identical for both columns of each row, it will not affect on song quality. Even, some low magnitude extra values can be added anywhere over the sampled data of the song without hampering the originality of the song [1, 2, 5]. Now, if we add some secret information or code over total sampled data range in some selected rows intelligently, secret message with song signal may also be transmitted without affecting the quality of original song. Let, **x(n,2)** is set of total sampled data of a song signal and **p(n,2)** is extra sampled values after equaling all rows for both columns, i.e. $x(n,2) \cong x(n,2) + p(n,2)$, because they are different by very small values for each position.

The technique has been organized into three stages. In the first stage, adjustment of phase is performed by separating phase and amplitude signals from original and secret songs followed by coinciding both signals at the similar point of time. In second stage, representing secret song is achieved with set of lower amplitude values and embedding the same in original signal. Finally, fabricating of secret code is done in third stage.

The proposed encoding technique has been outlined in subsection 2.1 that of decoding is in subsection 2.2 which shown in the figure 1.

2.1. Embedding

Coinciding original and amplitude songs at the similar point of time has been represented in the subsection 2.1.1. The method of secret song embedding has been depicted in the subsection 2.1.2 whereas authentication technique is described at subsection 2.1.3.

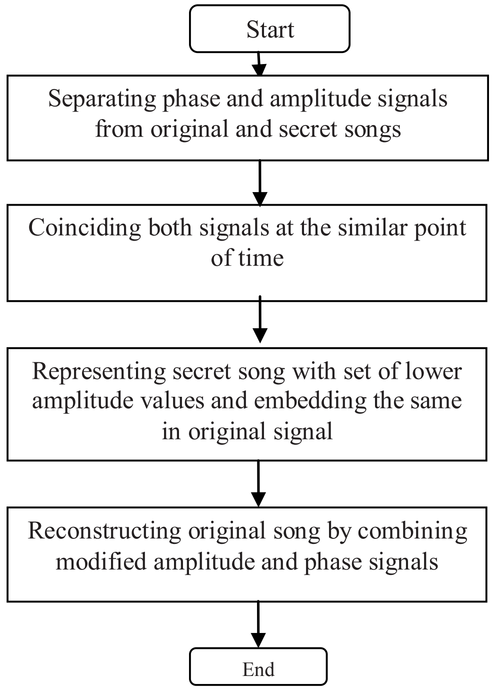

Fig 1. Block diagram of proposed technique

2.1.1. Adjustment of Phase

Original and secret signals may be different in phase at the similar point of time, correlating between the phases, the shifting of one is needed to the similar position of other phase.

The method of phase shifting of secret song's coinciding with the phase of original song is described as follows.

Input: Original and secret song signals.
Output: Modified song with adjustment of phase value.
Method: Separating amplitude and phase of original song and coinciding phase of secret and original signals are described in the following steps.

Step 1: Separate phase and amplitude signals from original and secret songs as shown in figure 2 and 3.
Step 2: If both phase signals do not coincide at the similar point of time, shift the phase signal of secret song along time line and shifted position are counted. Shifting of phase is shown in figure 4.
Step 3: Reconstruct the original and secret songs by combining phase and amplitude components respectively as reverse way of step 1.

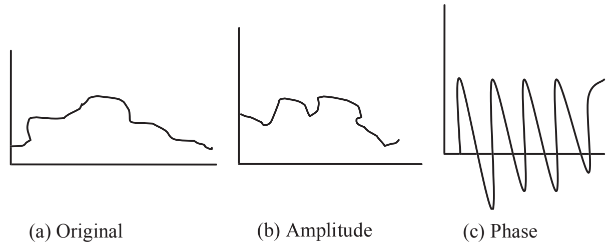

Fig 2: Separating amplitude and phase from original

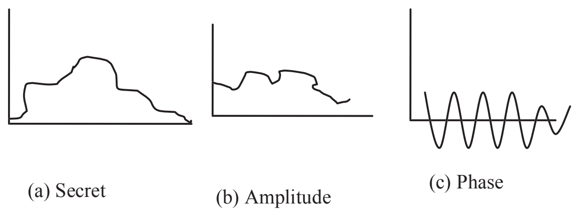

Fig 3: Separating amplitude and phase from secret song

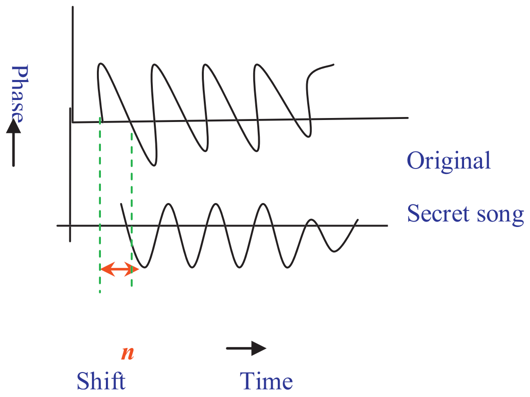

Fig 4: Sifting of phase

2.1.2. Amplitude Encoding

To embed the amplitude values of secret song over the original song signal need to convert secret song amplitude values into lower values as after embedding of these extra values will not alter the song quality. Therefore, amplitude encoding is necessary for embedding into original song signal as an authenticate code. The amplitude encoding is described in the following steps.

Input: Original and secret song signals.

Output: Modified song with embedding secret song.

Method: Separating amplitude and phase of original song and embedding secret song with original song signal are described in the following steps.

Step 1: Stretching the amplitude signal of secret song (along time line) to the $4/5^{th}$ times of original signal by dividing each amplitude value by n [a number , where $n*L1=(4/5)*L2$, L1= length of secret song and L2 = length of cover song] and generating lower amplitude values will be embedded with the cover song as secure code.

Step 2: Lower amplitude values is added with the original song signal as follows:

*i. Find amplitude value from 1^{st} position of secret song and divide it by n [same as step 1]. Say, the resultant value is "**val**".*
ii. Make amplitude values of two channels equal at the n^{th} position of original song signal.
*iii. Add equal amplitude value (**val**) with the any of two channels at n^{th} position.*
Iv Repeat the step i to iii for adding all remaining lower magnitude values (original amplitude values of secret song divide by n) with an alternative channel position at next n^{th} position for each value set.

Step 3: Reconstruct the original song applying reverse process.

For secret song of stereo type, information of two channels are to be equally distributed throughout the cover signal as described in above steps.

2.1.3. Fabricating Secret Code

This is one of the additional measures appended with signal to attain an extra security level. The shifted value of hidden song's phase [step 2 of 2.1.1], amplitude values difference of n^{th} position of all value sets [step 2(iii) of 2.1.2] to be added with consecutive positions into higher frequencies (above 20,000 Hz) in alternative manner with both channels, i.e., total secret parameters will be distributed between two channels in a predefined gap.

Therefore, if any value changes during processing, it will create a difference with the n^{th} position of amplitude value of previous set of consecutive values. As the intensity of extra added amplitude values of secret song is small, will not affect overall song's quality.

2.2. Authentication

Extracting the secret song from original and justifying the phase shift, it is easily detect the original song from similar collection of same. The extraction procedure is also very simple and described through following steps.

Input: Modified song with embedding secret song.

Output: Original song signal.

Method: Generating original song from modified song is described in the following steps.

Step 1: Remove the authenticated code from its respective positions of sampled values of amplitude of cover signal from all n^{th} positions. Finding extra added value by comparing values of both channels at n^{th} position and deduct the same from the cover signal.

Step 2: The amplitude value of first position of secret song will be multiply of n of the extra value remove from n^{th} position.

Step 3: Find the next extra added value by considering next n^{th} position value of alternative channel of signal, remove same from song signal.

Step 4: Continue step 3 up to the last available row position of added extra value over cover signal.

Step 5: Remove all appended extra values from higher frequencies (above 20,000 Hz) from cover signal.

If the secret song properly extracted from original signal, i.e., it has not altered during processing.

3. Experimental Results

Encoding and decoding techniques are applied over 1 minute recorded song and results in each intermediate step have been outlined in sections 3.1.1 to 3.1.4. The results are discussed in two sections out of which 3.1 deals with result associated with IPRSESS and that of 3.2 gives comparative analysis with existing techniques.

3.1. Results

For experimental observation strip of one minute rock song ('100 Miles From Memphis, sung by Sheryl Crow) has been taken. The sampled value of the song is given in table 1 as a two dimensional matrix. Figure 5 shows amplitude-time graph of the original signal. IPRSESS is applied on this signal and output generated in the process is shown in figure 7 (number of sampled values is 2646000). Figure 8 shows the difference of amplitude values of original and modified song after embedding hidden song and authenticated code. From figure 8 it is seen that the deviation is very less which will not affect the quality of the song at all.

3.1.1. Original recorded song signal (1 minute)

The values for sampled data array x(n,2) from the original song is given in table 1. The graphical representation of the original song, considering all rows (2646000) of x(n,2) is given in the figure 5.

Table 1: sampled data array X(N,2)

Sl no	x(k,1)	x(k,2)
…	…	…
	0	0.0001
	0.0000	0.0000
	-0.0009	-0.0009
	-0.0006	-0.0007
	-0.0014	-0.0014
	-0.0016	-0.0017
	-0.0023	-0.0022
	-0.0027	-0.0027
	-0.0022	-0.0021
…	…	…

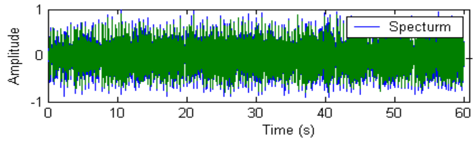

Fig. 5 Original song (‘100 Miles From Memphis’, sung by Sheryl Crow)

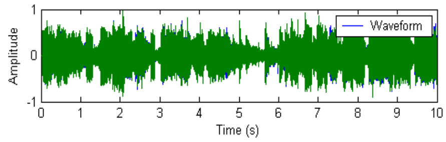

Fig. 6 Hidden song (‘Awakening of cheerful feelings on arriving in the country’, sung by Sir Simon Rattle)

3.1.2. Hidden song (10 sec)

The graphical representation of the concealed song is shown in the figure 6.

3.1.3. Modified song after adding hidden song and authenticated code (1 minute)

The graphical representation of the composite song signal is shown in the figure 7.

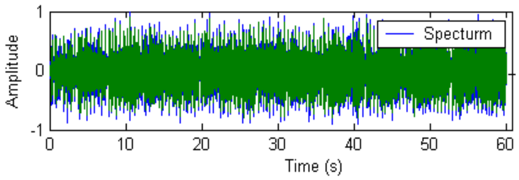

Fig. 7 Modified song after adding hidden song and authenticated code

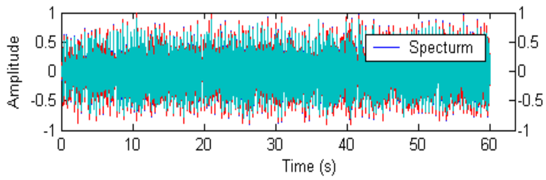

Fig 8: The difference of sample values between signals (figure 4 and figure 6)

3.1.4. Difference of amplitude values between modified and original songs (1 minute)

The graphical representation of the amplitude difference of original and modified songs is shown in the figure 8.

4. Analysis

Various algorithms are available for embedding information with audio signals [6, 7]. They usually do not care about the quality of audio but we are enforcing our authentication technique by keeping the quality of song intact. A comparative study of properties of IPRSESS method with data hiding via phase manipulation of audio signals (DHPMA)[3] before and after embedding secret message/modifying parts of signal (16-bit stereo audio signals sampled at 44.1 kHz.) is given in table 2, 3 and 4. Average absolute difference (AD) is used as the dissimilarity measure between original and modified song to justify the modified song. Whereas a lower value of AD signifies lesser error in the modified song. Normalized average absolute difference (NAD) is quantization error is to measure normalized distance to a range between 0 and 1. Mean square error (MSE) is the cumulative squared error between the embedded song and the original song. A lower value of MSE signifies lesser error in the embedded song. The SNR is used to measure how much a signal has been tainted by noise. It represents embedding errors between original and modified song and calculated as the ratio of signal power (original song) to the noise power corrupting the signal. A ratio higher than 1:1 indicates more signal than noise. The PSNR is often used to assess the quality measurement between the original and a modified song. The higher the PSNR represents the better the quality of the modified song. The quantitative estimation for the quality of extracted modified song with reference to the original song can be expressed as normalized cross correlation gives maximum value of (NC) as unity. From the Table 2-4 in proposed technique, the measuring value of NAD equal (0.0201) signifies lesser error in the modified song signal. The MSE has been obtained a lowest average of the squared difference between the intensity of the original and modified song at each amplitude location equal (1.0029e-005) with ideal value in the NMSE equal (2.5441e+003). The SNR getting a high-quality ratio equal (34.0553dB) means that the lesser obtrusive of the embedding errors in the modified song. The PSNR is getting a higher of PSNR equal (49.8148dB), a logically a higher value of PSNR is high quality modified song. Thus from our experimental values of benchmarking parameters (NAD, MSE, NMSE, SNR and PSNR) in proposed method are expressed overall better performance without changing the audio quality of song.

The embedding technique has been applied over 5 different types of audio signals (cover songs). Song 1 and song 2 are rock songs, song 4 is pop song, song 5 is classic song and song 3 is Rabindra Tagore's song. It was embedded the five different types of 16-bit stereo audio signals sampled at 44.1 kHz. Table 3 gives the experimental results in terms of SNR (Signal to Noise Ratio) and PSNR (Peak signal to Noise Ratio) which shows better performances in IPRSESS.

Table 4 represents comparative values of Normalized Cross-Correlation (NC) and Correlation Quality (QC) of proposed algorithm with DHPMA. The Table 5 shows PSNR, SNR, BER (Bit Error Rate) and MOS (Mean opinion score) values for the proposed algorithm. Here all the BER values are 0 which again indicate good performance for IPRSESS.

Figure 9 summarizes the results of this experimental test. It shows that performance of IPRSESS is stable for different types of audio signals. This quality rating (Mean opinion score) is computed by using the formula (1).

$$Quality = \frac{5}{1 + N * SNR} \tag{1}$$

Where N is a normalization constant and SNR is the measured signal to noise ratio. The ITU-R Rec. 500 quality rating is perfectly suited for this task, as it gives a quality rating on a scale of 1 to 5 [11]. Table 6 shows the rating scale, along with the quality level being represented.

Table2 : MD, AD, NAD, MSE and NMSE

Sl No	Statistical parameters for differential distortion	Value using IPRSESS	Value using DHPMA
1	MD	0.0262	3.6621e-004
2	AD	0.0024	2.0886e-005
3	NAD	0.0201	0.0063
4	MSE	1.0029e-005	1.4671e-009
5	NMSE	2.5441e+003	8.4137e-005

Table3 : SNR and PSNR

Sl No	Statistical parameters for differential distortion	Value using IPRSESS	Value using DHPMA
1	Signal to Noise Ratio (SNR)	34.0553	40.7501
2	Peak Signal to Noise Ratio (PSNR)	49.8148	45.4226

Table 4 : NC and QC

Sl No	Statistical parameters for differential distortion	Value using IPRSESS	Value using DHPMA
1	Normalised Cross-Correlation (NC)	1	1
2	Correlation Quality (QC)	-0.1144	-0.5038

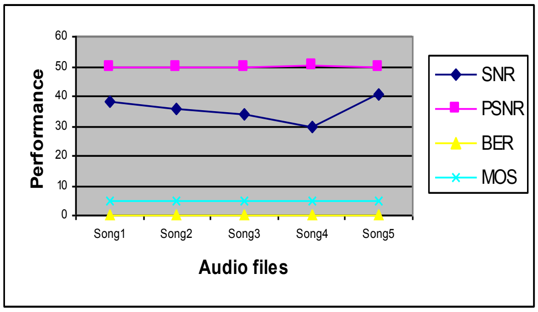

Fig 9: Performance for different audio signals

Table 5 : SNR, PSNR BER, AND MOS for different songs

Audio (1 minute)	SNR	PSNR	BER	MOS
Song1	38.2069	49.9868	0	5
Song2	35.6781	49.9718	0	5
Song3	34.0553	49.8148	0	5
Song4	29.4515	50.0430	0	5
Song5	40.3602	49.9499	0	5

Table 6 : ITU-R Rec. 500 quality rating

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

5. Conclusion and Future works

In this paper, an algorithm to hide a strip of audio over song signal has been proposed which will use to authenticate song signal with the help of embedding secret code as well as secret signal. The hidden song and message will not affect the song quality but it will ensure to detect the distortion of the signal. Additionally, the proposed algorithm is also very easy to implement.

This technique is developed based on the observation of characteristics of different songs but the mathematical model for representing the variation of those characteristics after modification may be formulated in future. The technique can be extended to embed an image into an audio signal instead of text and audio. The perfect estimation of percentage of threshold numbers of sample data of song that can be allowed to change for normal conditions will be done in future with all possibilities of errors.

References

[1] Uttam Kr. Mondal and J.K. Mandal, "A Practical Approach of Embedding Secret Key to Authenticate Tagore Songs(ESKATS)," in Proc. Wireless Information Networks & Business Information System Proceedings (WINBIS'10), ISSN 2091-0266,organized by Rural Nepal Technical Academy (Pvt.) Ltd , Nepal, Volume 6, pp. 67-74, Number 1, 2010.

[2] Uttam Kr. Mondal and J.K. Mandal , " A Novel Technique to Protect Piracy of Quality Songs through Amplitude Manipulation (PPAM)," presented at the International Symposium on Electronic System Design (ISED 2010), ISBN 978-0-7695-4294-2 , pp 246-250, 2010.

[3] Xiaoxiao Dong , Mark F. Bocko and Zeljko Ignjatovic , "Data Hiding Via Phase Manipulation of Audio Signals," in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004) , ISBN 0-7803-8484-9 ,Vol 5, pp. 377-380, 2004.

[4] G. Erten, F. Salam,"Voice Output Extraction by Signal Separation," in Proc. ISCAS '98 ,ISBN 07803-4455-3,Vol 3, pp. 5 - 8, 1998.

[5] Uttam Kr. Mondal and J.K. Mandal, "Song Authentication Technique Through Concealment of Secret Song (SATCon)," presented at IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011, MIT, Anna University, Chennai., pp. 145-150, June 3-5, 2011.

[6] F.A.P. Petitcolas, R.J. Anderson and M.G. Kuhn,"Information Hiding –A Survey," in Proc. the IEEE, 87(7), pp. 1062-1078, 1999.

[7] R.J. Anderson, and F.A.P. Petitcolas, "On the Limits of Steganography," in Proc.IEEE Journal of Selected Areas in Communications, 16(4), pp. 474-481, 1998.

[8] George R. Doddington," Speaker Recognition- Identifying People by their Voices," in Proc of the IEEE, VOL 73, NO. 11, pp 1651-1665, Nov 1985.
[9] H. Suzuki, H. Zen, Y. Nunkuku, C. Miyajima, K. Tokuda, and I. Kitumuru,," in Proc. Speech Recognition Using Voice- Characteristic Dependent Acoustic Models", ICASSP 2003,Vol 3, pp 740-743, 2003.
[10] Kun S. Lin and Gene A. Frantz," in Proc. Voice Characteristics Conversion", IEEE Transactions on Consumer Electronics, Vol. CE-30, No. 4, pp 598-603, November 1984.
[11] M. Arnold, "Audio watermarking: Features, applications and algorithms", IEEE International Conference on Multimediaand Expo 2000.